A binary complex of the catalytic subunit of cAMP-dependent protein kinase and adenosine further defines conformational flexibility

Narendra Narayana[1]*, Sarah Cox[1†], Nguyen-huu Xuong[2], Lynn F Ten Eyck[1,3], and Susan S Taylor[1,3]

Background: cAMP-dependent protein kinase (cAPK), a ubiquitous protein in eukaryotic cells, is one of the simplest members of the protein kinase family. It was the first protein kinase to be crystallized and continues to serve as a biochemical and structural prototype for this family of enzymes. To further understand the conformational changes that occur in different liganded and unliganded states of cAPK, the catalytic subunit of cAPK was crystallized in the absence of peptide inhibitor.

Results: The crystal structure of the catalytic subunit of mouse recombinant cAPK (rC) complexed with adenosine was solved at 2.6 Å resolution and refined to a crystallographic R factor of 21.9% with good stereochemical parameters. This is the first structure of the rC subunit that lacks a bound inhibitor or substrate peptide. The structure was solved by molecular replacement and comprises two lobes (large and small) which contain a number of conserved loops.

Conclusions: The binary complex of rC and adenosine adopts an 'intermediate' conformation relative to the previously described 'closed' and 'open' conformations of other rC complexes. Based on a comparison of these structures, the induced fit that is necessary for catalysis and closing of the active-site cleft appears to be confined to the small lobe, as in the absence of the peptide the conformation of the large lobe, including the peptide-docking surface, does not change. Three specific components contribute to the closing of the cleft: rotation of the small lobe; movement of the C-terminal tail; and closing of the so-called glycine-rich loop. There is no induced fit in the large lobe to accommodate the peptide and the closing of the cleft. A portion of the C-terminal tail, residues 315–334, serves as a gate for the entry or exit of the nucleotide into the hydrophobic active-site cleft.

Addresses: [1]Department of Chemistry and Biochemistry and School of Medicine, University of California, San Diego, La Jolla, CA 92093-0359, USA, [2]Departments of Biology, Physics, Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0359, USA and [3]San Diego Supercomputer Center, La Jolla, CA 92186, USA.

[†]Present address: DuPont Merck Pharmaceutical Company Experimental Station, E336/36A, Rt 141 and Henry Clay Road, Wilmington, DE 19880-0361, USA.

*Corresponding author.
E-mail: Narendra@chem.ucsd.edu

Key words: conformational changes, crystal structure, electrostatic complementarity, phosphorylation, protein kinase

Received: 10 March 1997
Revisions requested: 21 April 1997
Revisions received: 16 June 1997
Accepted: 17 June 1997

Structure 15 July 1997, 5:921–935
http://biomednet.com/elecref/0969212600500921

Introduction

The protein kinases are a large and diverse family of enzymes that are critical for many signal transduction pathways in eukaryotic cells [1]. cAMP-dependent protein kinase (cAPK), one of the smallest and simplest protein kinases, serves as a template for the entire family because all share a conserved catalytic core [2]. The structure of the catalytic subunit (C) of cAPK revealed a bilobal structure with ATP binding at the base of the cleft between the two lobes and peptide binding at the edge of the cleft [3,4]. A set of 11–12 essentially invariant residues that are spread throughout the conserved core are positioned in an identical way in all of the protein kinases. Although some of these residues appear to play a purely structural role, most cluster around the active-site cleft and contribute either to nucleotide binding or phosphoryl transfer.

To understand the role of these conserved residues and to appreciate the extensive network of interactions that permeates the catalytic core requires a variety of approaches that include fluorescence, kinetic and mutational analysis [5,6]. These solution methods are essential. Crystallography, on the other hand, provides a view of fixed conformations. The numerous binary and ternary complexes of the C subunit solved so far provide evidence for both 'closed' and 'open' conformations [2–4,7–13], and demonstrate the flexibility of the enzyme. Kinetic analysis of the reaction pathway shows the importance of this flexibility. Binding of substrates induces a conformational change that leads to a closing of the active-site cleft. Phosphoryl transfer is very fast (500/sec) while the slower (20/sec) rate-limiting step is product release, which also includes the conformational changes that lead to the opening of the cleft [14,15].

We describe here a binary complex of the mouse recombinant catalytic (rC) subunit and adenosine (rC·Ade). This is the first structure of the C subunit that has been solved with no bound peptide. The structure reveals that the conformation of the ledge on the large lobe, on which the peptide docks, does not change as a consequence of substrate binding; it is a static docking surface. The substrate-induced conformational changes necessary for catalysis are thus entirely a consequence of movements in the small lobe. The structure described here represents an intermediate conformation, between the previously described closed and open conformations, and allows us to dissect more carefully the specific movements within the subdomains of the small lobe that enable the cleft to close or open. The abbreviations used to represent the complexes discussed here, and their available accession numbers, are given: rC–PKI(5–24) (rC·I, 1apm); rC–PKI(5–24)–Mg·ATP (rC·I·ATP, 1atp); mC–I_2·PKI(5–24) (mC·I, 1ctp, 1cmk); C–PKI(5–24)–Mn·AMP-PNP (mC·I·AMP-PNP, 1cdk); rC–PKI(5–24)–adenosine (rC·I·Ade); rC–adenosine (rC·Ade); rC–PKS(5–24)–ADP (rC·Sp·ADP); and rC–PKS(5–24)* (rC–Sp*) (also see Table 1).

Results

Description of the enzyme structure

The different phosphorylation states of the C subunit correspond to multiple isozymes. Isoform II of rC has 350 amino acids (M_r ~40,000) with 3 autophosphorylation sites, Ser10, Thr197 (termed here Thp197) and Ser338 (termed Sep338) [16]. In the final refined protein atomic coordinates, residues 1–11 are missing due to disorder. The protein molecule consists of a smaller N-terminal domain, having predominantly β structure, and a larger C-terminal domain comprised largely of α helices (Figure 1), similar to the previously described rC structures [2]. The enzyme is in an intermediate conformational state, as compared to the closed and open conformations previously solved (see sections on ligand-induced conformational changes and the active-site cleft for details) [8,9,12].

Different types of β turns are found in the regions of chain reversal and many of the invariant residues that contribute to enzyme function are located in these loops. Table 1 summarizes the different loops with their associated functions and mean B factors. The structure, in the context of these conserved and variable loops, is described below in some detail. The extended chain that links the nonconserved A helix to the first structural element of the small lobe, strand β1, is anchored firmly to the small lobe. Strands β1 and β2 form a β hairpin with the turn at residues 53 and 54. The glycine-rich sequence (Gly50-Thr-Gly-Ser-Phe-Gly55) of this loop is a highly conserved feature of all protein kinases and is referred to as the 'glycine-rich loop'. A type IV bend at residues 62–65 links β2 to β3; this segment, β1–β3, functions as a highly conserved nucleotide-positioning motif. β1–β3 is then connected to the helical segment in the small lobe containing helices B and C. The C helix occupies a critical place at the active site cleft which will be discussed later. A long loop (eight residues) containing a type I β turn at residues

Table 1

Loops in the structure of the rC·Ade binary complex.

Loop	Residue numbers	Type of turn	Average B factor (Å^2)†	Function
A→β1	32–42	VII	53, 58	Anchors C terminus
β1→β2	53–54	I	85, 50	Part of glycine-rich loop, anchors α-, γ- phosphates‡
β2→β3,B*	63–66	IV	61, 56	–
B→C	83	I	88, 34	–
C→β4	98–105	I	42, 27	Multiple contacts with large lobe and the A helix
β4→β5	113–114	γ	56, 36	
β5→D	121–127	I	25, 29	Linker‡
D→E, β6*	136–139	VII	23, 26	Hydrophobic
β6→β7	164–171	I	25, 24	Catalytic loop‡
β7→β8	176–178	IV	37, 43	Anchors C-terminal tail
β8→β9	184–187	I	26, 31	Mg^{2+} positioning loop‡
β9→EF	193–206	VII, inverse γ	37, 27	Activation segment‡
EF→F	212–216	irregular	57, 42	Interacts with the catalytic loop
F→G	234–242	irregular	48, 30	P–11 pocket
G→H	253–262	I	37, 54	–
H→I	274–287	inverse γ, II, VII	36, 48	–
I→J	294–301	III	43, 50	–
J→Phe350	308–350	extended, II	58, 53	Covers adenosine-binding pocket

The β strands and α helices are labeled sequentially as β1, β2 etcetera and A helix, B helix and so on, respectively. The arrow denotes a loop between the secondary structure elements. The 3_{10} helices are also considered part of loops. *The secondary structures are contiguous. The turns of type I, II, III, IV, and VII refer to β turns [58,59], and the γ turns were based on the classifications of Millner-White *et al*., [60]. †The two numbers refer to rC·Ade and rC·I·Ade complexes, respectively, the B factors were averaged over mainchain atoms. ‡Indicates the presence of a conserved loop in the structure.

100–103 connects the C helix with strand β4. This loop is packed firmly against the large lobe. Strands β4 and β5 form another β hairpin, where the loop is variable. The two β hairpins and the middle strand β3 form an antiparallel twisted β sheet. Strand β5 is followed by a linker region (121–127) which contains two highly conserved glycine residues and is the only mainchain entry point into the large lobe in the conserved core. Except for the B helix the secondary structure elements of the small lobe are conserved in all other protein kinase structures reported to date [17–23]. The conserved loops in the large lobe have been described previously [24] and are defined structurally in Table 1.

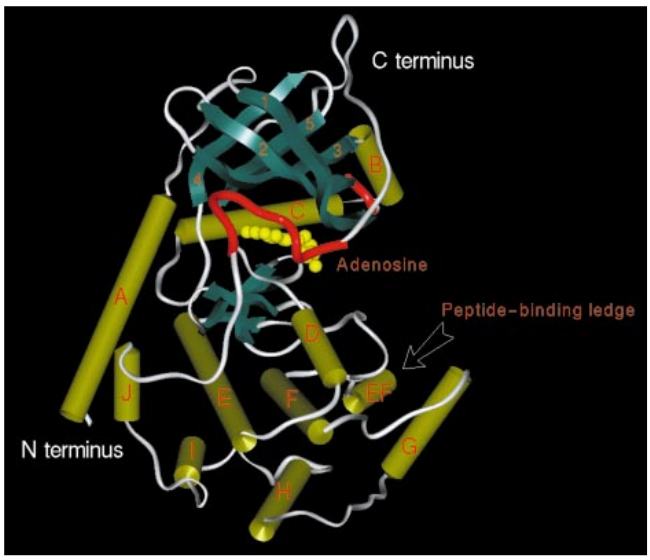
The nucleoside-binding site

Adenosine is bound to a distinctive protein kinase nucleotide-binding fold, which includes the glycine-rich loop, with interactions essentially similar to those observed in the ternary complex rC·I·Ade [12]. In this binary structure, however, unlike in the ternary complexes, the ribose portion of the adenosine is solvent accessible due to the partial opening of the cleft by the glycine-rich loop and the absence of shielding by the peptide inhibitor (Figure 1). The adenine ring fits snugly into the hydrophobic pocket at the junction of the two lobes. The purine base forms three hydrogen bonds, as shown in Figure 2. The ribose adopts a major C3′-endo pucker and assumes an *anti* ($\chi = -143°$) conformation with respect to the adenine base. The exocyclic hydroxyl O3′ of the ribose donates a hydrogen bond to the carbonyl of Glu170 (3.1 Å), but the 2′-OH is too far from the carboxylate of Glu127 to form a hydrogen bond as it does in the rC·I·Ade ternary complex. In the ternary complex Glu127 interacts with the O2′ in the ribose, the P–3 Arg in PKI(5–24) (see Figure 3 legend for explanation of peptide inhibitor nomenclature), and Tyr330 in the C-terminal tail via a water molecule [12]. There is no evidence for this ordered water molecule in the binary structure. Thus in the absence of the peptide this coordination is not achieved. The torsion angle C3′–C4′–C5′–O5′ corresponds to a different rotomer from that found in the ternary complex of rC, peptide inhibitor and Mg·ATP (rC·I·ATP; PDB code, 1atp) [4].

The triphosphate/metal ion binding subsite

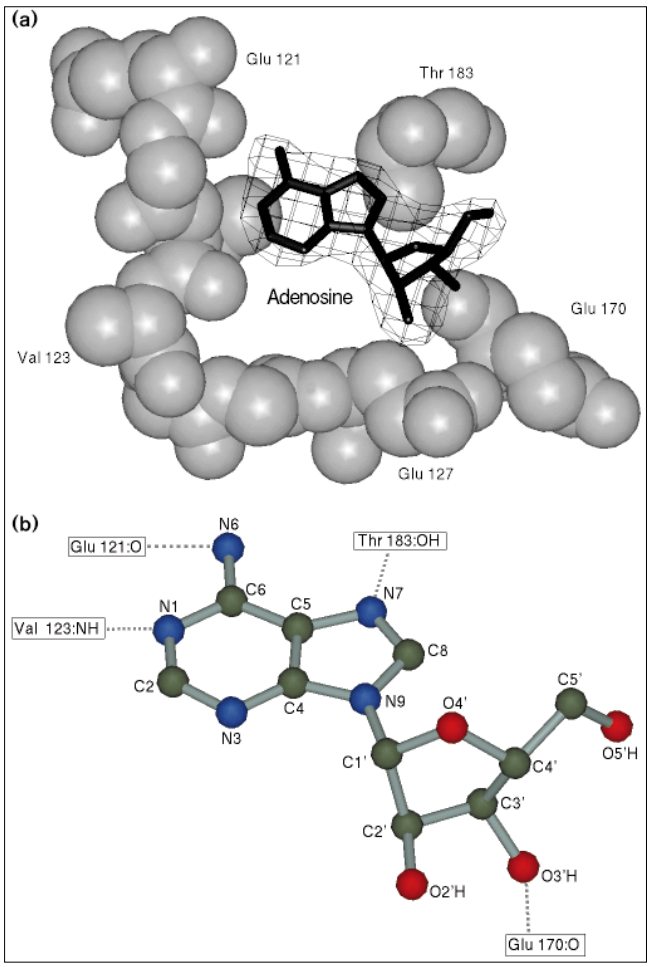
Several attempts to crystallize an rC·ATP complex were unsuccessful (NN, unpublished results). This may in part be due to the slow ATPase activity of the enzyme. The divalent metal ions in cAPK are involved in interactions with the phosphates of ATP, and are known to serve multiple roles in enzyme catalysis and product dissociation [25]. In the rC·Ade binary complex the triphosphate group is missing; this space is exclusively filled by solvent. In order to understand the binding mode of the ATP co-factor, the missing triphosphate group was modeled into this subsite using the coordinates of rC·I·ATP[4]. Except for the van der Waals' interactions between the triphosphate moiety and the glycine-rich loop, in the binary complex the triphosphate and the divalent metal ions can maintain interactions with the enzyme similar to those in rC·I·ATP [4]. The enzyme residues require only minor conformational changes in the sidechains (Lys72, Glu127, Asn171 and Asp184) to form stable interactions with the modeled phosphate groups. In the absence of the triphosphate group and the peptide inhibitor, the glycine-rich loop is shifted away from the large lobe. This results in fewer van der Waals interactions with the modeled triphosphate group. The torsional angle C3′–C4′–C5′–O5′ is significantly different in the ATP molecule (74°) [4] than compared to that in the adenosine containing complexes (–38°, rC·I·Ade; and –59°, rC·Ade). This rotation about the C4′–C5′ bond is necessitated by the presence of the phosphate groups, which are involved in stable interactions with the metal ion ligand

Figure 1



The architecture of rC with bound adenosine. The α helices and β strands are shown as cylinders and ribbons, respectively. The connecting loops, including 3_{10} helices, are shown as tubes. The secondary structure elements and the residues they encompass are as follows: α helices, A (13–31), B (76–82), C (84–97), D (128–135), E (140–160), EF (207–211), F (217–233), G (243–252), H (263–273), I (288–293), J (302–307); 3_{10} helices, a (40–42), b (169–171) c (185–187), d (202–204), e (277–279), f (295–297); β strands, 1 (43–52), 2 (55–62), 3 (67–75), 4 (106–112), 5 (115–120), 6 (161–163), 7 (172–175), 8 (179–183) and 9 (188–192). The EF helix was considered as part of a loop preceding the F helix in the earlier publications. The red ribbons represent the highly mobile section of the C-terminal tail (residues 318–326) and the tip of the glycine-rich loop (residues 52–55). The loops connecting the secondary structure elements are listed in Table 1. Adenosine is shown in ball-and-stick representation in yellow. Abbreviations relevant to the structure are: small lobe (40–120); linker (121–127); large lobe (128–300); C-terminal tail (301–350); small lobe anchor (301–314); C-terminal gate (315–334); C-terminal acidic cluster (328–334); and C-terminal anchor (335–350). The α helices A, B, I and J are not conserved. (The figure was made using the graphics program InsightII [61].)

Figure 2



The adenosine-binding pocket. **(a)** The difference Fourier map is contoured at six standard deviations and was computed using the Fourier terms $(F_o–F_c)e^{i\alpha}$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively, and α is the calculated phase using the final TNT refined model without adenosine. The refined adenosine molecule is superposed on the electron density. **(b)** Schematic diagram showing the hydrogen-bonding interactions between adenosine and the enzyme residues. Dashed lines indicate hydrogen bonds; atoms are shown in standard colours. (The figure was made using the graphics program Insightll [61].)

clusters and the enzyme, and is required to prevent stereochemical clashes with the enzyme.

The peptide-binding site

In all the previously determined structures of rC and the myristylated catalytic subunit (mC) [2–4,7–13], with the exception of the low-resolution apo mC structure [8], the enzyme is in a complex with the peptide inhibitor, PKI (5–24), or a substrate analogue of PKI(5–24) [11,12,26–28]. As summarized in Figure 3, in the absence of the petide inhibitor the extensive docking surface is unchanged even though the contact surface involves several loops that link elements of secondary structure. Loops that specifically contribute to peptide recognition, including the catalytic and the P+1 loops, do not change. The P+1 leucine sits in a hydrophobic pocket comprised of the sidechains of Leu198, Pro202 and Leu205. The P–2, P–3 and P–6 arginine residues of the peptide inhibitor make ionic interactions with the glutamate residues 170 and 230, 127, and 203, respectively. Three of these residues are in loops: Glu127 in the linker, Glu170 in the catalytic loop and Glu203 in the P+1 loop. The P–11 phenylalanine of the inhibitor is located in a hydrophobic pocket consisting of Arg133, Tyr235, Pro236 and Phe239. A stretch of residues, 235–239, links helices F and G. In essence the peptide-binding region is in place and ready to receive the peptide without the need for any conformational changes within the large lobe of the enzyme (Figures 3c and 3d). The preference for P–2, P–3 and P–6 arginines is explained by the complementary negatively charged surface at the active-site cleft.

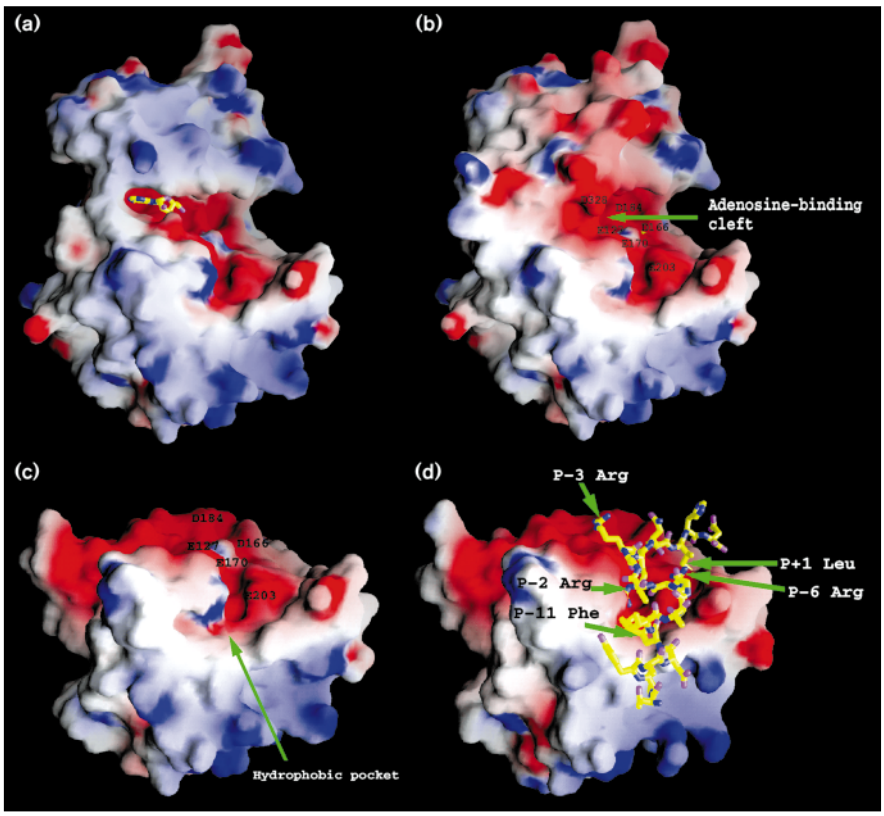
The active-site cleft

The wedge-shaped active-site cleft is located between the upper (small) and lower (large) domains. The slender edge is near the linker region (residues 121–127) which is termed the base of the cleft. The outer thicker edge is positioned towards the residues Lys83 in the small lobe and Leu198 in the large lobe and can be described as the mouth of the cleft. The upper and the lower surfaces of the active-site cleft are lined by the residues from the small and large lobes, respectively. One side of the cleft is masked by the C helix in the small lobe. In the standard view shown in Figures 3a and 3b, this C helix forms the 'back' side of the cleft. The nucleoside is bound deep within the cleft; the outer edge of the cleft is the binding site for the peptide inhibitor (Figure 1) [7,11].

Interdomain spacing

In this discussion we use three distinct parts of the molecule and their relative positions to define the overall conformation of the molecule. Three regions of the molecule and their juxtaposition can be said to describe the degree of 'openness' of the structure (S Shaltiel *et al.*, unpublished data): the interaction between His87 and Thp197 at the mouth of the cleft interface, including the relative orientation of the small lobe with respect to the large lobe; the position and degree of disorder of the glycine-rich loop at the top of the cleft; and the C-terminal tail which covers the front of the cleft (Table 2). Prior to this study, all the rC complexes studied had exhibited a 'closed' conformation [12]. The native mC and the binary mC·I complex crystallized in an 'open' form [8], whereas the mC·I·ATP and mC·I·AMP-PNP complexes crystallized in the closed conformation [8,13]. The present structure adopts an overall conformation that is intermediate. For our discussions on domain orientation, the representative examples chosen are the binary mC·I complex for the open conformation [8], the ternary rC·I·ATP complex for the closed conformation

Figure 3



The electrostatic potential surface around rC displayed at the level of the solvent-accessible surface. The electrostatic potential was computed at neutral pH using the program GRASP [62]. The charge distribution is color-coded with blue for positive (≥ 10 kT/e), red for negative (≤ –10 kT/e) and white for neutral, where k is the Boltzmann constant, T is the temperature, and e is the electronic charge. **(a)** Standard front view of the rC·Ade complex without the gate (residues 315–334). The active-site cleft is comprised of several acidic residues resulting in a negatively charged surface. Adenosine is shown in stick form and colored in cyan. **(b)** The same view as in (a) with the gate shown; adenosine is covered by the gate on the front side of the active-site cleft. **(c)** Electrostatic surface of the large lobe of the rC·Ade complex in the same view as in (a) and (b). After the calculation of the electrostatic potential using the whole rC molecule, the surface for the large lobe only was displayed. There is no change in the peptide-docking surface. **(d)** Similar view as in (c) with bound peptide inhibitor PKI(5–24) (shown in yellow stick formation) in the ternary rC·I·Ade complex [12]. PKI(5–24) corresponds to residues 5–24 (TTYADFIASGRTGRRNAIHD) in the heat-stable protein kinase inhibitor PKI [63]. PKS(5–24) represents the substrate peptide where residues N20 and A21 in PKI(5–24) are changed to AS. PKS(5–24)* represents the same peptide phosphorylated at Ser21. The letter P denotes the Ser21 phosphorylation site in the substrate peptide, sites N-terminal and C-terminal to this are represented as P–1 and P+1 etcetera. The P–2, P–3 and P–6 arginine residues of the inhibitor PKI(5–24) are bound to the negatively charged surface on the large lobe.

[4], and the present structure for an intermediate conformation. The radii of gyration in the open, intermediate and closed conformations are 21.0, 20.3 and 20.0 Å, respectively. In comparison to the closed conformation, the small lobe in the present structure is rotated by about 4° relative to the large lobe (Figure 4); this angle is 11° in the open form. See Table 2 for details on some characteristics of domain motions.

Interdomain interactions

The C helix, comprised of residues 85–95 in the small lobe, fills the back surface of the active-site cleft and is

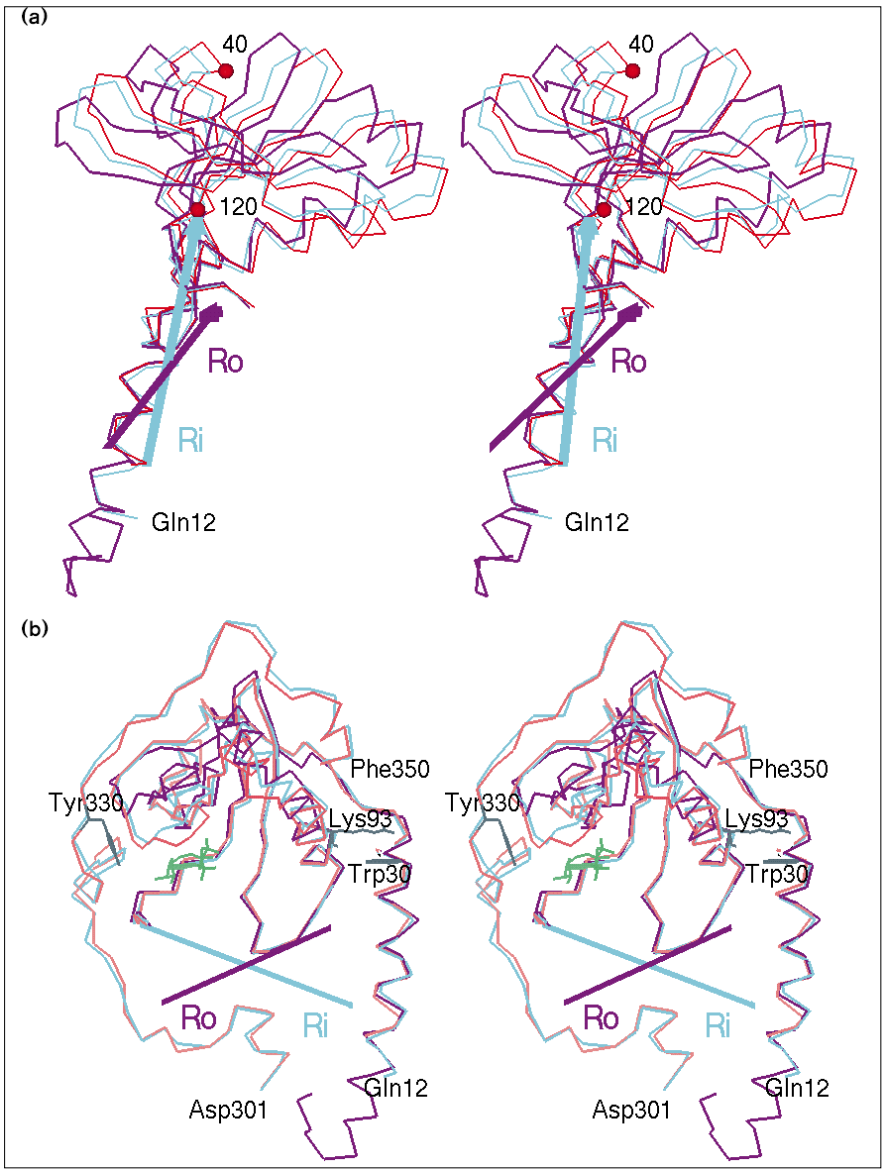
Table 2

Some parameters describing global conformation.

	mC·I (open)	rC·Ade (intermediate)	rC·I·ATP (closed)	rC·Sp·ADP (closed)	rC·Sp* (closed)
Radius of gyration (R_g) (Å)†	21.0	20.3	20.0	20.0	19.8
Distance between His87 and Thp197 (Å)	7.1	3.2	2.7	2.9	2.7
Distance between Ser53 and Gly186 (Å)	14.2	11.8	10.4	8.9	10.7
Distance between Glu170 and Tyr330 (Å)	14.6	8.6	8.2	8.1	8.3
C-terminal segment (318–326)	not seen	less ordered	ordered	ordered	ordered
Average B factor (Å^2)	–	85	47	43	48

Secondary structure content (48%; α helices, 3_{10} helices and β strands) remains unaltered during conformational changes from open → intermediate → closed forms. †R_g was calculated using all non-hydrogen atoms of rC. Inclusion of substrate peptide or peptide inhibitor increased the value of R_g by ≤ 0.1 Å. The atoms involved in distance measurements were Nε2 of His87, phosphate oxygen of Thp197, Cα atoms of Ser53, Gly186, O of Glu170 and the sidechain hydroxyl of Tyr330.

Figure 4



Superposition of the open, intermediate and closed conformations. **(a)** Stereo view of the small lobes. The Cα atoms of the large lobe (128–300) in mC·I (open) and rC·Ade (intermediate) were superposed onto the equivalent atoms in rC·I·ATP (closed). The root mean square deviation (rmsd) between mC·I and rC·I·ATP was 0.42 Å and that between rC·Ade and rC·I·ATP was 0.51 Å. The open mC·I, the intermediate rC·Ade and the closed rC·I·ATP are colored in purple, cyan and red, respectively. The small lobe in both mC·I and rC·Ade exhibit concerted shifts relative to that seen in rC·I·ATP. Ro represents the rotation axis about which the small lobe (excluding residues 49–57) of the open structure is rotated by 11° to achieve the position seen in the closed form. Similarly, the small lobe of the intermediate structure is rotated by 4° about the axis denoted by Ri. **(b)** Stereo view approximately perpendicular to that shown in (a). The C-terminal tails of rC·Ade and rC·I·ATP are also displayed. ATP is shown in green, Trp30, Arg93 and Tyr330 are in black. The two axes Ro and Ri cross over at about 46°. Portions of the small lobe that are less influenced by the closing or opening of the cleft are seen in this view. (The figure was made using the graphics program Insightll [61]).

involved in hydrogen bonding and numerous van der Waals' interactions with residues in the large lobe, as well as with residues in the nonconserved N-terminal and C-terminal segments that flank the conserved core. His87 is within hydrogen-bonding distance of Thp197 and in van der Waals' contact with Arg165, both in the large lobe. There are also van der Waals' interactions between the C helix and the Asp184-Phe185-Gly186 (DFG) loop in the large lobe. In addition to these interactions with the large lobe, the C helix interacts (via Arg93) with Trp30 in the nonconserved A helix, which spans both lobes of the protein kinase core [29,30]. Finally, the C helix makes an electrostatic contact with the C-terminal carboxylate which is buried in the small lobe. Thus residues in the C helix play a prominent role in holding different segments of the enzyme, widely separated in the amino acid sequence, in a folded, active conformation. The C helix also includes the conserved residue Glu91 which interacts with another conserved residue Lys72 in the formation of an ion pair. This ion pair links the two major conserved elements of secondary structure in the small lobe, the β sheet and the C helix. The Lys72, in turn, bridges the α- and β-phosphates of ATP. Figure 5 shows how key residues from different segments of the enzyme interact with the C helix. Both the glycine-rich loop (via Phe54) and the C helix help to position the DFG loop and thus the γ-phosphate of ATP. The C helix in the present structure resembles the conformation it assumes in the closed conformation.

Ligand-induced conformational changes

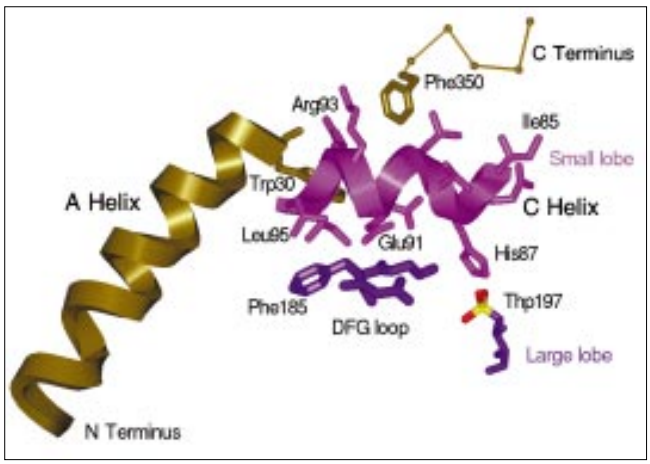
The rC subunit has three flexible segments: the first 10–15 residues at the N terminus; the glycine-rich loop; and a section of the C-terminal tail (315–334) which crosses the active-site cleft from the large lobe to the small lobe. The flexibility of these segments is highly dependent on what ligands are bound. Specifically, both the glycine-rich loop and the C-terminal segment are well ordered in the rC·I and rC·I·ATP complexes [3,4,13]. As discussed later, a segment of the C-terminal tail (315–334) appears to function as a gate for nucleoside/nucleotide entry or exit.

The C-terminal domain is involved in extensive intradomain hydrogen bonding and hydrophobic interactions which contribute to the stability of the core. Least-squares superposition of the Cα atoms of the large lobe (residues 128–300) in the different rC and mC complexes gives a root mean square deviation (rmsd) in the range 0.5–1.0 Å. There are no systematic shifts and no major deviations in any segment of the superposed large lobes. Therefore, the origin of the observed rmsd of about 0.5 Å may be attributed to errors in the atomic coordinates, the varying resolution of the structures, and true minor conformational changes in the different complexes. In rC the small N-terminal lobe has about twice as many hydrogen bonds between the secondary structure elements as does the large C-terminal lobe. During the transition from open→intermediate→closed conformational state, these hydrogen bond interactions in the β sheet are preserved, allowing for the concerted motions of the secondary structure elements in the smaller domain. This structural setting favors cooperative shifts in the polypeptide segments without much expense of energy. The short extended peptide segment, residues 121–127 including the conserved glycine residues 125 and 126, which connects the two domains has the capability of hinge bending, as described below.

Analysis of the transition from open→intermediate→closed forms

Some aspects of domain movements between the open and closed conformational states were described previously [10]. Comparison of the different structures, with respect to closure or opening of the active-site cleft, is appropriately done by superposition of the equivalent residues in the conserved large lobe [10]. The relative orientation of the A helix (residues 15–31) with respect to the large lobe is fixed in the open, intermediate, and closed conformational states. The large lobes of the open (mC·I) and intermediate (rC·Ade) conformations were optimally superimposed on the large lobe of the closed ternary complex (rC·I·ATP) by rigid-body transformations using the algorithm developed by Diamond [31]. In this superposition the small lobes in the mC·I and the rC·Ade structures are rotated by 11° and 4°, respectively, relative to the small lobe in rC·I·ATP (as shown in Figure 4).

Figure 5



The environment of the C helix. Residues in the C helix interdigitate with residues in the large lobe. In addition, the C helix residues (85–95) are involved in important interactions with different segments of the enzyme which are widely separated in amino acid sequence. The conserved residue Glu91 interacts with another conserved residue, Lys72, which in turn ion pairs with the α- and β-phosphates of ATP. Thus, the C helix plays a central role in folding the protein into an active conformation as well as in enzyme catalysis.

The positions of the rotation axes for the open→closed and intermediate→closed transitions were calculated from data produced by Diamond's program [31]. The two axes pass within 0.5 Å of each other but cross at an angle of 46°. The superpositions of the small lobe Cα traces (residues 40–120, excluding the glycine-rich loop residues 47–59) have an rmsd within 0.6–0.7 Å. The residues that are likely to be involved in the hinge-bending motions (based on the least-squares fit of Cα atoms of the large lobes in open, intermediate, and closed forms) are: residues 30–40 between the A helix and the small lobe; the linker region (residues 121–127); and the ends of the glycine-rich loop. The rotation axes do not pass directly through any residues. The hinge-bending movement is accomplished by concerted changes of mainchain conformational angles throughout the linker region.

Crystal packing

The molecules in the crystal of the rC·Ade binary complex are arrayed in chains of molecules related by crystallographic 2_1 axes along the unit cell axis directions (a, b and c). There are 33 contacts less than 4 Å apart and five hydrogen bonds between adjacent molecules along the 2_1 axis parallel to b; and there are 31 close contacts and five hydrogen bonds between adjacent molecules along the 2_1 axis parallel to c. There are, however, no close contacts between adjacent molecules along the 2_1 axis parallel to a. Therefore, the molecules are primarily held together by intermolecular interactions along the b and c directions. There is one dominant contact region between the enzyme

molecules along the b direction, where residues 133, 134, 190, 192 and 194 are in close proximity of the symmetry-related residues 196, 241 and 328. This was also the binding region for residues P–8 (serine), P–9 (alanine) and P–10 (isoleucine) of the peptide inhibitor PKI(5–24) in the structures of binary and ternary complexes [7,12]. The NH1 and NH2 of Arg190 hydrogen bonds with Oδ2 of Asp241. Nε of Arg194 interacts with Oδ1 and Oδ2 of 328, and NH1 of 194 hydrogen bonds with Oδ2 of Asp328. The consecutive arginine residues 133 and 134 in the D helix stack hydrophobically against the nearby symmetry-related Trp196. This interaction was noted previously in different crystal forms of both rC and mC complexes [10,12], suggesting an important role for Arg133, Arg134 and Trp196 in protein–protein interactions. Trp196 is a major determinant for the recognition of the regulatory (R) subunit of cAPK, but not for PKI(5–24) [32] (RM Gibson, *et al.*, personal communication). Arg133, on the other hand is a major determinant for the recognition of PKI(5–24), but not for RIα, the type I regulatory subunit of cAPK [33]. These two surfaces thus define two distinct sites that are essential for high affinity binding of the two classes of physiological inhibitors of cAPK.

Surface residues from both the small and the large lobes contribute to the protein-protein interactions along the c direction. In this direction, charged residues in the small lobe and the C-terminal tail (Asp41, Asp44, Lys63, Glu334 and Arg336) are in contact with the residues of the symmetry-related large lobe (Arg137, Pro141, Arg144, Ser259, His260, Phe261, Ser262, 263, Thr299 and Glu311): the carboxylate oxygens Oδ1 and Oδ2 of Asp44 are hydrogen bonded with the amide and the hydroxyl groups of the symmetry-related Ser263; the Oε2 of Glu311 is involved in an ionic interaction with NH2 of Arg336; and the carboxylate oxygens Oε1 and Oε2 of Glu334 accept hydrogen bonds from the hydroxyl group of Thr299 and NH1 of Arg144, respectively. The short contiguous stretch (residues 259–263) involved in these intermolecular contacts is in the loop preceding the I helix in the large lobe; the electron densities for these interacting residues are well defined. Except for the intermolecular triad, Arg133, Arg134 and Trp196 interaction, all other intersubunit interactions in the present complex without a peptide inhibitor are different from those seen in the previously solved structures of rC complexes [12].

Discussion

Domain movements

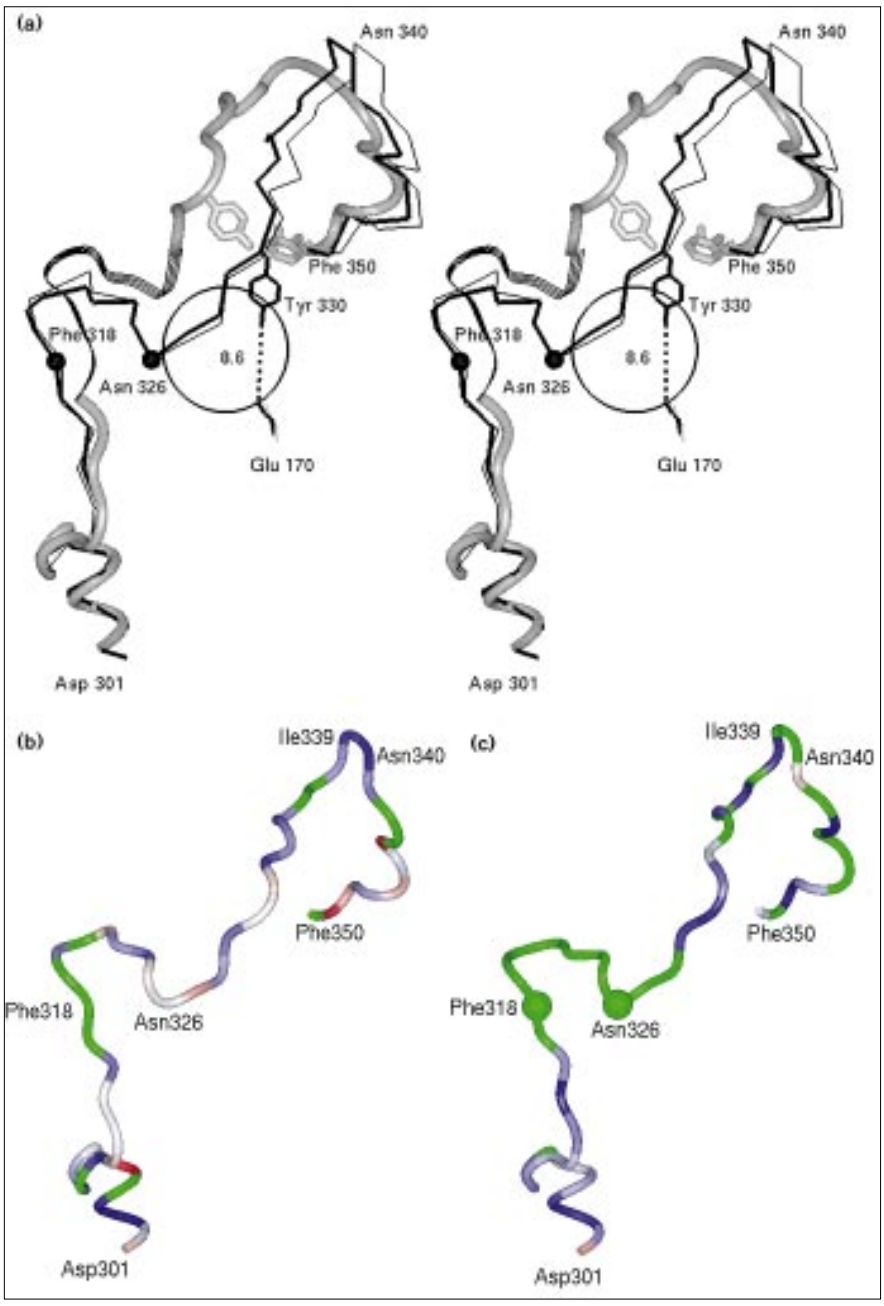
The relative position of domains in crystal structures of liganded and unliganded proteins provide striking examples of protein flexibility. Crystallographic evidence for such conformational changes involving domain motions in small molecule kinases were first shown for hexokinase [34] and have been described in great detail for adenylate kinase [35]. The mechanisms for domain movements in these proteins involve both hinge and shear motions [36]. The larger hinge motions typically involve changes in the main-chain torsion angles, while the smaller shear motions involve interdigitating sidechains with changes in the side-chain torsion angles of less than 15°.

Previous structures of the C subunit revealed both closed and open conformations [2–4,7–13]. With the exception of a binary complex of the mammalian enzyme with a modified peptide inhibitor [9], all assumed a closed conformation with the major variations being in the position of the glycine-rich loop. Whether this segment of the nucleotide-positioning motif is closed or open depends on which specific ligands occupy the active-site cleft. The fully closed loop is seen only in the ternary complex containing rC, Mg^{2+}·ATP and PKI(5–24). The overall conformation of the enzyme in the adenosine binary complex described here assumes an intermediate state that is distinct from both the open and closed forms seen previously. Least-squares superposition of the Cα atoms of the large lobe on the corresponding atoms in the closed rC·I·ATP complex indicates a rotation of the small lobe, including the glycine-rich loop, away from the large lobe. The small lobe is in a position that is intermediate between its position in the open (mC·I) and the closed (rC·I·ATP) complexes. Rigid-body transformations [31] indicated that the small lobe is rotated by 4° relative to the small lobe in rC·I·ATP in contrast to 11° in the open mC·I structure. The distances between the Cα atoms of residue 53 at the tip of the glycine-rich loop and residue 186 in the stable loop of the large lobe in mC·I, rC·Ade and rC·I·ATP are 14.2, 11.8 and 10.4 Å, respectively. The position of this flexible loop is not influenced by crystal contacts. The distance between His87 and Thp197 in the open, intermediate and closed conformations is 7.1, 3.2 and 2.7 Å, respectively.

Qualitatively the domain motion of the conserved kinase core can be described in terms of three bodies: the large lobe, the small lobe minus the glycine-rich loop, and the glycine-rich loop. There is not a simple single hinge point as suggested by Olah *et al.*, [37]. The large lobe and the small lobe are related by one rotation axis; within the small lobe the position of the glycine-rich loop is determined by a second rotation axis. In the transition from the closed to the intermediate conformation the small lobe rotates by 4° while the glycine-rich loop rotates by 10°. In the transition from the closed to the open conformation the small lobe rotates by 11° but the glycine-rich loop rotation is still 10°. In other words, the glycine-rich loop has already rotated maximally in the intermediate structure. The additional movement is achieved by rotation around the large lobe/small lobe axis.

The fact that each pair of structures can be superposed by a single rotation to reposition the small lobe would seem to suggest a simple hinge motion. However, the finding

Figure 6



The C-terminal gate. **(a)** Stereo view of the Cα atoms of the C-terminal tail in the open mC·I (shown as a tube), intermediate rC·Ade (thick black line), and the closed rC·I·ATP (thin black line) structures after superposition of equivalent Cα atoms in the large lobe (128–300). Residues 318–326, that are missing in the 1ctp coordinates of mC·I, are observed in the 1cmk PDB entry. This segment is shown as a ribbon. The circle represents the position of the adenosine-binding site. The C-terminal gate, comprised of residues 315–334 in rC·I·ATP and rC·Ade, shields the nucleoside on the front of the active-site cleft. The same view of the C-terminal tail in **(b)** the rC·I·ATP and **(c)** rC·Ade structures. The residues are color-coded according to their B factors. B values in the range 0 to 25 Å^2, 25 to 50 Å^2, and greater than 50 Å^2 are color-ramped red to white, white to blue, and green, respectively. The C-terminal gate is ordered in the closed rC·I·ATP conformation. In contrast, rC·Ade exhibits high B values implying flexibility as in the open conformation where this segment is disordered.

that the rotation axes for each pair are different and approximately intersect at a 46° angle shows conclusively that the actual transformation must be more complex than a simple hinge in the linker region (residues 121–127) as proposed by Olah *et al.*, [37]. Even if the differing lattice contacts in the open, intermediate and closed forms partly contribute to this 46° difference in rotation axes, the fact remains that the correlated conformational changes that bring about the transformations involve the same stretch of residues.

The position of the C-terminal tail is also diagnostic for open and closed conformations (S Shaltiel *et al.*, unpublished data). In the closed conformation the C-terminal segment (residues 315–334) covers the adenine-binding site on the front side of the active-site cleft (Figures 1 and 6a) and is stable as indicated by low and uniform B factors (Figure 6b). In the open conformation this stretch of residues was either not seen (1ctp) [9] or was shifted away from the cleft (1cmk) [8]. In the present binary complex this segment is seen in a conformation similar to that observed

in the closed conformation, however, the electron density for this segment was weak and exhibited higher B factors than the remaining C-terminal tail residues. Thus this segment is ordered in the closed rC·I, rC·I·ATP, and mC·I·AMP-PNP structures, less ordered in the present binary complex, and disordered in the open mC.I. complex. This contiguous stretch of residues (315–334) may therefore serve as a gate for nucleoside/nucleotide entry or exit. With the sliding of this flexible C-terminal gate and the concomitant movement of the flexible glycine-rich loop away from the large lobe, it is possible that the nucleoside/nucleotide can enter or exit the cleft when the enzyme is in an intermediate conformation. Further quantitation of the spatial separation of the domains, including the C-terminal tail, is provided for different conformational states in Table 2.

Based on the crystal structures of the mC and rC complexes, it is very likely that the concerted motion of peptide segments and the resilient characteristic of the glycine-rich loop in the small lobe, together with a less extensive interdomain surface, imparts to the enzyme the capability of closing and opening the active-site cleft. In solution, it is plausible that the enzyme is in a dynamic oscillating state between the open and closed conformations with the equilibrium shifted towards the open form in the absence of substrates or inhibitors. The presence of metal ions may also influence this equilibrium [25]. We believe for two reasons that the conformation of the rC·Ade complex in the crystal is essentially the predominant conformation in solution, remaining aware of the possibility that lattice interactions affect the protein conformation. Firstly, there are only two hydrogen bonds between the small lobe and the symmetry-related large lobe: the sidechain atoms of Asp44 in the small domain interact with the amide and the hydroxyl groups of the symmetry-related Ser263 in the large lobe. In addition, as noted previously, there are only a few residues in the small lobe and the C-terminal tail (Asp41, Asp44, Lys63, Glu334 and Arg336) within 4 Å of the symmetry-related atoms. Secondly, the crystallization conditions used here are similar to those used in the crystallisation of the rC·I·Ade ternary complex [12]. The present binary complex rC·Ade, however, crystallized in a different crystal form. Therefore, in the present complex the influence of the crystallization conditions and the lattice contacts on the relative orientation of the small lobe with respect to the large lobe may be marginal. Presumably, the rC·ATP conformation in the crystal or in solution would be similar to that seen in the present rC·Ade complex. As discussed previously, the modeled triphosphate maintains interactions with the key residues, such as Lys72 and Glu91 and other residues in the active-site cleft, although there are fewer van der Waals' interactions with the glycine-rich loop. The binding of the peptide inhibitor or the substrate peptide to rC results in bringing the glycine-rich loop close to the triphosphate or diphosphate moiety, thus causing a complete or partial closure of the cleft, as seen in rC·I·Ade, rC·I·ATP and the complex of rC with PKS(5–24) and ADP (rC·Sp·ADP). So far, the fully closed conformation of this loop is seen only in rC·I·ATP.

The peptide-binding ledge

The binary structure of the C subunit described here reveals for the first time a conformational state of the enzyme without a bound peptide inhibitor. In the absence of the peptide inhibitor, PKI(5–24), there appears to be no significant changes in the enzyme residues on the surface of the conserved large lobe where the peptide docks (Figures 1 and 3c) [12]. The absence of the inhibitor and the triphosphate of ATP does affect the relative position of the small lobe and the glycine-rich loop with respect to the large lobe, however, none of the residues and loops in the large lobe that interact, either directly or indirectly, with the peptide show any major change. The positions of the catalytic loop, the P+1 loop, and the DFG loop do not change, although each contact PKI(5–24) in binary and ternary complexes. This surface on the large lobe thus serves as a very stable ledge on which the peptide docks. No induced conformational changes occur in the large lobe as a consequence of peptide binding.

Ligand binding in the active site

One of the striking features revealed by the different rC and mC complexes is the varying interdomain spacing. If one views the active-site pocket as a wedge, the adenine ring is bound at the apex of the wedge and this portion of the cleft changes very little due to opening and closing. This is in contrast to the edge of the wedge, where large changes are seen. The C helix which contacts the large lobe as well as both the N-terminal and C-terminal segments of the enzyme plays a prominent structural role by linking the apex of the active-site cleft to its outer edge. In the closed conformation the edge is firmly held by the contact between His87 and Thp197.

The simplest form of molecular recognition is based on the size and shape of the binding site. The enzyme's inter-domain motion, as well as the intrinsic flexibility of the glycine-rich loop and the small lobe in general may be used advantageously for designing potent enzyme inhibitors with varying shapes and sizes. In particular, inhibitors could be designed with respect to the portion of the ligand that binds towards the thicker edge of the cleft. The adenine-binding pocket near the linker region appears to bind planar moieties with one or two fused cyclic rings [38–40] (NN, unpublished data). The adenine-binding pocket is less influenced by domain movements than the mouth of the cleft, as it is located near the hinges or the points of rotation of the domain. Therefore, substitution or modification of this portion of the ligand seems to be of limited use in the design of inhibitors.

Nonetheless, the purine-binding site can be affected by the conformational changes within the domains.

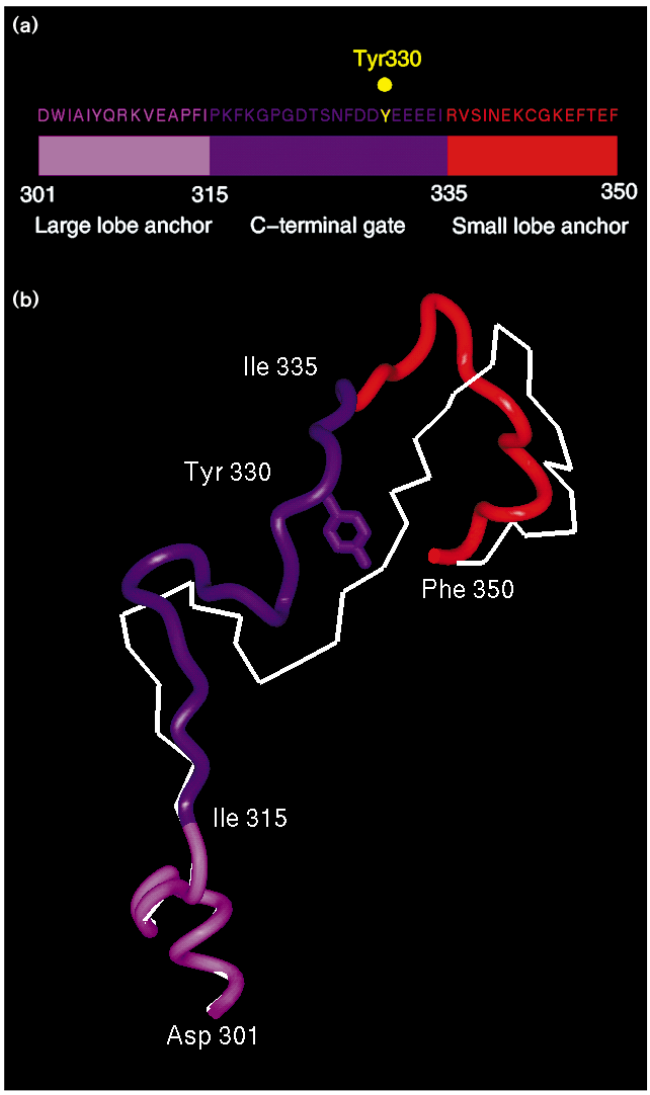
C-terminal tail

The C-terminal tail (residues 301–350) embodies three distinct segments (Figure 7): residues 300–314 serve to anchor the tail to the large lobe (these residues have low B factors in all C subunit structures solved to date); residues 315–334 serve a 'gating' function that allows for access to the nucleotide-binding site; and residues 335–350 provide a stable anchor to the small lobe and also have low B factors. The degree of order in the gating segment varies widely depending on whether the enzyme is in a closed, intermediate, or open conformation. This segment can be divided into two subcomponents based on the variability of the B factors: residues 328–334 and residues 318–326. Residues Asp328-Asp-Tyr-Glu-Glu-Glu-Glu334 typically have low B factors and in the closed conformation this segment serves as a latch for the gate, with Tyr330 making contact with both the nucleotide and the P–3 arginine in PKI(5–24). The carboxylate groups also make numerous contacts with the small lobe in the closed conformation. In the open conformation this segment is well ordered but exposed to solvent. This difference was also reflected in the reactivity of the carboxylate groups with a water soluble carbodiimide: in the free enzyme these groups are reactive but in the ternary complex they are protected [41]. The segment comprised of residues 318–326 is quite different; this section was either not observed or was poorly ordered in a very different conformation in the open state [8,9]. In the intermediate conformation this segment is closed but the B factors are high. The flexible character of the C-terminal gate together with the flexibility of the nearby glycine-rich loop may facilitate access of the nucleotide into or out of the active-site cleft.

Thermal stability

The negatively charged phosphate groups of the nucleotide are surrounded by a number of conserved negatively charged residues: Glu91, Asp166 and Asp184. In the presence of divalent cations, the phosphate anions and the nearby enzyme residues interact favourably due to balanced electrostatic interactions. In the absence of cations, however, the nucleotide binds poorly presumably due to charge repulsion. The thermostability of the free enzyme is less than that of the rC·Mg^{2+}·ATP complex (F Herberg, personal communication). The greater heat stability of the rC·Mg^{2+}·ATP complex is due to salt bridges involving ATP, Mg^{2+} and the active-site residues. The thermal stability of the enzyme in the presence of adenosine is similar to the stability of the rC·Mg^{2+}·ATP complex and metal ions are not required for the enhanced thermostability. In the absence of both triphosphate and metal ions the crevice is filled with polar solvent molecules [12] which may suffice to create a stable electrostatic environment.

Figure 7



The C-terminal tail. (a) Schematic picture of the three distinct segments of the C-terminal tail (residues 301–350): the large lobe anchor, the C-terminal gate, and the small lobe anchor. Tyr330 (shown in yellow) makes contact with both the nucleotide and the P–3 arginine in PKI(5–24). (b) The three sections of the C-terminal tail are shown in pink, purple and red in the open mC·I structure; the closed ternary rC·I·ATP is shown in white.

Biological implications

The protein kinases are a large family of enzymes with two main classes—enzymes catalyzing the phosphorylation of serine/threonine residues, and those catalyzing tyrosine phosphorylation. Although this family is very diverse, all protein kinases share a conserved catalytic core. One of the simplest protein kinases is cAMP-dependent protein kinase (cAPK), and the catalytic (C) subunit of this enzyme serves as a template for the catalytic core of all eukaryotic protein kinases [2]. To date,

the C subunit of cAPK is the only protein kinase that has been crystallized in an active form in the presence of both substrates and inhibitors.

We report here the structure of the catalytic subunit of mouse recombinant cAPK (rC) in complex with adenosine. The overall structure shows similarities to the previously determined structures of rC complexes, comprising a bilobal structure containing a number of conserved loops. The adenosine binary complex assumes an 'intermediate' conformation relative to the 'open' and 'closed' conformations of the enzyme described previously [2–4, 7–13]. A comparison of various closed, open and intermediate conformations shows that the only conserved loop at the active-site cleft to move during the transition between these states is the so-called glycine-rich loop, which lies over the phosphate groups of bound ATP. This study sheds some light on the ligand-induced conformational changes. The intrinsic flexibility of the glycine-rich loop seen in this and previously solved structures of rC complexes, as well as in other serine/threonine- and tyrosine-specific protein kinases, suggest that it is a genuine phenomena of the protein kinase structure. Conformational flexibility of this loop is important for catalysis [42]. In some cases this loop may also inhibit catalysis, as in calcium/calmodulin-dependent protein kinase I [43], and the cyclin A–Cdk2–p27 inhibitory complex [44]. The mobility of the glycine-rich loop and the nearby nonconserved C-terminal gate provide an easy route to the active site for nucleotides. This mobility, together with hinge-bending motions of the small lobe, contribute to an 'induced fit' mechanism for catalysis.

The description presented here of a binary complex of rC with adenosine is the first structure of a cAPK C subunit to be determined with no bound peptide. The structure reveals that the ledge on which the peptide docks is very stable. Although the characteristic negative electrostatic surface of this ledge at the active-site cleft strongly influences substrate recognition, no conformational changes are induced as a consequence of peptide binding.

Materials and methods

Crystallization

The mouse rC was expressed in *Escherichia coli* and purified as reported previously [16]. Adenosine was purchased from Sigma Chemical Co. The rC·Ade binary complex crystals were grown at 4°C by the hanging-drop vapour diffusion method. The crystallization droplet contained protein (0.5 mM), adenosine (3 mM) and 2-methyl-2,4-pentanediol (MPD; 4%) in 100 mM bicine buffer at pH 8.0. The crystallization well solution was made up of 15% MPD and 100 mM ammonium sulfate in bicine buffer (100 mM; pH 8.0). The crystals grew over the course of 8–12 weeks to a final size of $0.2\times0.3\times0.5$ mm^3. Crystallization trials using polyethylene glycol 400 as precipitant, in accordance with the protocol described in the rC·I and rC·I·ATP complexes [45], did not yield crystals.

Table 3

Molecular replacement results.

Search model*	mC·I (1ctp)
Resolution (Å)	10–4
Reflections (>3σ)	2 830
Rotation search	
P1 cell for the probe model (Å)	120×120×150 (cartesian cell)
Patterson vector length (Å)	45
Sampling interval for θ_2 (°)	2.5
Rotation peak before/after	θ_1 = 73.286/65.972
PC refinement (Euler angles; °)	θ_2 = 25.000/25.614
	θ_3 = 73.286/82.915
Peak ranking	14/1 (cc = 0.134)†
Translation search	
Sampling interval (Å)	1.0
Centroid of the molecule	x = 0.42, y = 0.37, z = 0.04 (fractional coordinates)
Translation peak height	18σ
R factor (%)	41.0

*Residues 32, 34, 39, 44, 65, 69, 108, 124 and 348 were appropriately modified to match the rC sequence. Enzyme residues without the myristic acid portion were used in the search model. Residues 1–6, 318–326, 339, 340 and some sidechains were not seen in the enzyme structure. †Correlation coefficient; cc of the next highest peak is ~0.06 and close to the noise level.

X-ray data collection

The crystal was exposed to a graphite monochromated Cu Kα beam from a Rigaku RU-200 rotating-anode X-ray generator operating at 5 kW power. X-ray diffraction data to 2.6 Å resolution were collected at 4°C from a single crystal with Xuong-Hamlin multiwire area detectors [46]. The crystal has the symmetry of space group $P2_12_12_1$ with the unit cell dimensions a = 55.85 Å, b = 73.23 Å and c = 98.72 Å; and one binary complex molecule per asymmetric unit (V_m = 2.5 Å^3/Da; 51% solvent content) [47]. The structure-factor amplitudes (F_{obs}) were obtained from the raw data using the UCSD area detector data processing programs [48]. The X-ray data is 86% complete with an R_{sym} ($\Sigma|I_{obs}-I_{ave}|/\Sigma\ I_{ave}$) of 0.094. The Wilson B is 45 Å^2 for 2.6 Å data. The crystal belongs to a new orthorhombic crystal form not seen previously for either the mammalian C subunit (mC) or the recombinant rC complexes [2–4,7–13].

Structure solution and refinement

The structure of mouse rC complexed with adenosine was determined by the molecular replacement (MR) method using the Patterson correlation (PC) refinement option in the X-PLOR program [49,50]. The atomic coordinates of both the closed and open conformations of rC (1apm) [7] and mC (PDB code, 1ctp) [9], respectively, were used in the rotation function calculations. Several search models were made using these coordinates – all enzyme residues, the small lobe and the large lobe. Computations were performed with and without sidechains in the models, at different resolution ranges. An unambiguous and conspicuous rotation function solution after PC refinement was determined using the coordinates of mC. Translation search using the highest PC refined rotation function peak resulted in a clear MR solution (18σ) with an initial R factor of 41.0% for data between 10 and 4.0 Å resolution. Initial crystal packing analysis at this stage indicated that the residues Arg133 and Arg134 were involved in hydrophobic interactions with the symmetry-related Trp196, as seen previously in C subunit complexes [10]. This MR solution exhibited favorable intermolecular interactions and there was no overlap between symmetry-related molecules. Details of the MR solution are listed in Table 3.

Table 4

Refinement results.

X-PLOR*	
Enzyme molecule	7–317, 327–338, 341–350
Small lobe	7–127, 307–317, 327–338, 341–350
Large lobe	128–306
(within small lobe)	
α helices (three groups)	15–30, 76–82, 84–97
β strands	43–48, 57–63, 67–75, 106–111, 115–120
Interconnecting segments	31–41, 98–105
Reflections ($F > 2\sigma_F$)	7 564
Resolution (Å)	10–2.7
R factor (%)†	23
TNT	
Reflections ($F > 0\sigma_F$)	10 720
No. of atoms	
in residues 12–350	2 801
in adenosine	19
in water oxygens	12
total nonhydrogen atoms	2 832
Resolution (Å)	10–2.6
R factor (%)	21.9
Stereochemical quality (rmsd)	
bond length (Å)	0.02
angles (°)	1.7
general/trigonal, planes (Å)	0.02 / 0.01
Mean B factors‡	
mainchain (Å^2)	41.4
sidechains (Å^2)	48.1
water oxygens (Å^2)	44.9

*For each of the rigid bodies listed the residues used in the X-PLOR calculations are shown. †R factor is defined as $100 \times \Sigma |F_o - F_c| / \Sigma F_o$. ‡B factors were defined with restraints.

The MR model was divided into two rigid domains: the small and the large lobe. The conformation of the large lobe is known to be strongly conserved in all presently known kinase structures. The small lobe was further broken down into five segments based on their secondary structures shown in Table 4. This segmented model of the enzyme was subjected to rigid-body refinement, a scheme similar to that adopted by Karlsson and co-workers [9]. The model was refined to an R factor of 39% with gradual addition of data to 2.7 Å resolution. Subsequent refinement continued according to simulated annealing slow-cool protocol with the temperature changing from 3000K to 277K in steps of 25K. This was followed by Powell minimization, overall B-factor and restrained individual B-factor refinements, which converged at R = 23.0% for 10–2.7 Å resolution ($F > 2\sigma$).

The X-PLOR refined model was further subjected to least-squares refinement using the conjugate direction algorithm in the TNT program [51]. A difference Fourier map calculated at this stage using 10–2.7 Å data clearly showed electron density for adenosine (Figure 2a). Refinement proceeded with the inclusion of adenosine and data to 2.6 Å resolution ($F > 0\sigma$). In order to have better estimates of the B values, the model was refined using restrained B and bulk-solvent correction options. The correctness of the model was corroborated by omit maps in which 10–15 residues surrounding adenosine, and also in the small lobe, were removed from the refinement and map calculations. Most of the sidechains and the missing residues (318–326 and 339–340) in the probe model were reliably built into the structure using the molecular graphics program package TOM-FRODO [52], via several rounds of difference Fourier maps and refinement. Computation of difference Fourier maps within the XtalView program suites [53] and the stereochemical check performed using the Biotech server [54,55] helped to locate the problem areas in the structure. It was possible to rebuild reliably the C-terminal gate residues only in the final rounds of refinement after resetting B values to 40 Å^2. The N-terminal residues (1–11) are invisible in the electron-density maps and are likely to be thermally disordered. A few ordered water molecules (12) were located in the $F_o - F_c$ maps.

Cross-validation calculations [56] were performed using a 'test' set of reflections (T) obtained by a 10% random selection from the observed reflections and a working set of reflections (W) containing the remaining 90% of reflections used in the refinement. The enzyme model with the adenosine and water molecules converged at R = 20.5% over set W and R_{free} = 34%. This final model gave an R factor of 21.9% for all data between 10 and 2.6 Å resolution ($F > 0\sigma$). The stereochemical quality of the structure and the backbone ϕ, ψ angles conform with the criteria of the program PROCHECK [54,57]. Of all nonglycine residues, 98% fall in either the most favorable or additionally allowed regions of the Ramachandran plot, and there are no nonglycine residues in the disallowed regions. (See Table 4 for some details on structure refinement).

Electrostatic surface potential

In order to calculate the electrostatic surface potential of rC, formal charges were assigned to one or two atoms of the residues that have a net charge. Assuming all lysines and arginines to have an overall +1e charge, all aspartates, glutamates and the C terminus to have an overall –1e charge, all histidines to be neutral, and the phosphorylated residues 197 and 338 to have an overall –2e charge, the net charge calculated for rC was –1e. The set parameters for the dielectric constant were 2 within and 80 outside the protein, and the solvent ionic strength was set at 0.0.

Accession numbers

Atomic coordinates of the structure have been submitted to the Brookhaven Protein Data Bank with accession code 1BKX.

Acknowledgements

We thank Victor Ashford, Chris Nielsen, Don Sullivan and Nick Nguyen for support at the NIH National Research Resource at University of California at San Diego multiwire area detector facility. NN thanks Bao Dang for assistance in growing crystals. We acknowledge the computing time from the San Diego Supercomputer Center for X-PLOR refinements. This work was supported by the Lucille P Markey Charitable Trust, by NIH grants GM19301 (SST), RR01644 and RR10748 (N-hX).

References

1. Hanks, S.K. & Quinn, A.M. (1991). Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* **200**, 38–62.
2. Knighton, D.R., *et al.*, & Sowadski, J.M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 407–414.
3. Zheng, J., *et al.*, & Sowadski, J.M. (1993). Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor. *Biochemistry* **32**, 2154–2161.
4. Zheng, J., *et al.*, & Sowadski, J.M. (1993). 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Cryst. D* **49**, 362–365.
5. Lew, J., Taylor, S.S. & Adams, J.A. (1997). Identification of a partially rate-determining step in the catalytic mechanism of cAMP-dependent protein kinase: a transient kinetic study using stopped-flow fluorescence spectroscopy. *Biochemistry* **36**, 6717–6724.
6. Grant, B.D., Tsigelny, I., Adams, J.A. & Taylor, S.S. (1996). Examination of an active-site electrostatic node in the cAMP-dependent protein kinase catalytic subunit. *Protein Sci.* **5**, 1316–1324.
7. Knighton, D.R., *et al.*, & Sowadski, J.M. (1993). 2.0 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with a peptide inhibitor and detergent. *Acta Cryst. D* **49**, 357–361.
8. Zheng, J., Knighton, D.R., Xuong, N.-h., Taylor, S.S., Sowadski, J.M. & Ten Eyck, L.F. (1993). Crystal structures of the myristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations. *Protein Sci.* **2**, 1559–1573.

9. Karlsson, R., Zheng, J., Xuong, N.-h., Taylor, S.S. & Sowadski, J.M. (1993). Structure of the mammalian catalytic subunit of cAMP-dependent protein kinase and an inhibitor peptide displays an open conformation. *Acta Cryst. D* 49, 381–388.
10. Karlsson, R., Madhusudan., Taylor, S.S. & Sowadski, J.M. (1994). Intermolecular contacts in various crystal forms related to the open and closed conformational states of the catalytic subunit of cAMP-dependent protein kinase. *Acta Cryst. D* 50, 657–662.
11. Madhusudan., *et al*., & Sowadski, J.M. (1994). cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphostransfer. *Protein Sci*. 3, 176–187.
12. Narayana, N., Cox, S., Shaltiel, S., Taylor, S.S. & Xuong, N.-h. (1997). Crystal structure of the poly-histidine tagged recombinant catalytic subunit of cAMP-dependent protein kinase complexed with the peptide inhibitor PKI(5–24) and adenosine. *Biochemistry* 36, 4438–4448.
13. Bossemeyer, D., Engh, R.A., Kinzel, V., Ponstingl, H. & Huber, R. (1993). Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn^{2+} adenylyl imidodiphosphate and inhibitor peptide PKI(5–24). *EMBO J.* 12, 849–859.
14. Grant, B.D. & Adams, J.A. (1996). Pre-steady-state kinetic analysis of cAMP-dependent protein kinase using rapid quench flow techniques. *Biochemistry* 35, 2022–2029.
15. Lew, J., Coruh, N., Tsigelny, I., Garrod, S. & Taylor, S.S. (1996). Synergistic binding of nucleotides and inhibitors to cAMP-dependent protein kinase examined by acrylodan fluorescence spectroscopy. *J. Biol. Chem*. 272, 1507–1513.
16. Herberg, F.W., Bell, S.M. & Taylor, S.S. (1993). Expression of the catalytic subunit of cAMP-dependent protein kinase in *Escherichia coli*: multiple isozymes reflect different phosphorylation states. *Protein Eng*. 6, 771–777.
17. De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O. & Kim, S.-H. (1993). Crystal structure of cyclin-dependent kinase 2. *Nature* 363, 595–602.
18. Zhang, F., Strand, A., Robbins, D., Cobb, M.H. & Goldsmith, E.J. (1994). Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. *Nature* 367, 704–711.
19. Hu, S.-H., Parker, M.W., Lei, J.Y., Wilce, M.C.J., Benian, G.M. & Kemp, B.E. (1994). Insights into autoregulation from the crystal structure of twichin kinase. *Nature* 369, 581–584.
20. Hubbard, S.R., Wei, L., Ellis, L. & Hendrickson, W.A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* 372, 746–754.
21. Owen, D.J., Noble, M.E.M., Garman, E.F., Papageorgiou, A.C. & Johnson, L.N. (1995). Two structures of the catalytic domain of phosphorylase kinase: an active protein kinase complexed with substrate analogue and product. *Structure* 3, 467–482.
22. Xu, R.-M., Carmel, G., Sweet, R.M., Kuret, J. & Cheng, X. (1995). Crystal structure of casein kinase-1, a phosphate-directed protein kinase. *EMBO J.* 14, 1015–1023.
23. Yamaguchi, H. & Hendrickson, W.A. (1996). Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* 384, 484–489.
24. Taylor, S.S., Knighton, D.R., Zheng, J., Ten Eyck, L.F. & Sowadski, J.M. (1992). Structural framework for the protein kinase family. *Annu. Rev. Cell Biol*. 8, 429–462.
25. Adams, J.A. & Taylor, S.S. (1993). Divalent metal ions influence catalysis and active-site accessibility in the cAMP-dependent protein kinase. *Protein Sci*. 2, 2177–2186.
26. Knighton, D.R., Zheng, J., Ten Eyck, L.F., Xuong, N.-h., Taylor, S.S. & Sowadski, J.M. (1991). Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* 253, 414–420.
27. Zetterqvist, O.Z., Ragnarsson, U. & Engstrom, L. (1990). Substrate specificity of cyclic AMP-dependent protein kinase. In *Peptides and Protein Phosphorylation* (Kemp, B.E. ed), pp. 171–187, CRC Press, Inc, Boca Raton, Florida, USA.
28. Walsh, D.A., Angelos, K.L., van Patten, S.M., Glass, D.B. & Garetto, L.P. (1990). In *Peptides and Protein Phosphorylation*. pp. 43–84 CRC Press, Inc., Boca Raton, Florida, USA.
29. Veron, M., Radzio-Andzelm, E., Tsigelny, I., Ten Eyck, L.F. & Taylor, S.S. (1993). A conserved helix motif complements the protein kinase core. *Proc. Natl. Acad. Sci. USA* 90, 10618–10622.
30. Herberg, F.W., Zimmermann, B., McGlone, M. & Taylor, S.S. (1997). Importance of the A-helix of the catalytic subunit of cAMP-dependent protein kinase for stability and orienting subdomains at the cleft interface. *Protein Sci*. 6, 1–11.
31. Diamond, R. (1992). On the multiple simultaneous superposition of molecular structures by rigid body transformations. *Protein Sci*. 1, 1279–1287.
32. Orellana, S.A. & McKnight, G.S. (1992). Mutations in the catalytic subunit of cAMP-dependent protein kinase result in unregulated biological activity. *Proc. Natl. Acad. Sci. USA* 89, 4726–4730.
33. Wen, W. & Taylor, S.S. (1994). High affinity binding of the heat-stable protein kinase inhibitor to the catalytic subunit of cAMP-dependent protein kinase is selectively abolished by mutation of Arg133. *J. Biol. Chem*. 269, 8423–8430.
34. Anderson, C.M., Zucker, F.H. & Steitz, T.A. (1979). Space-filling models of kinase clefts and conformation changes. *Science* 204, 375–380.
35. Gerstein, M., Schulz, G. & Chothia, C. (1993). Domain closure in adenylate kinase. *J. Mol. Biol*. 229, 494–501.
36. Gerstein, M., Lesk, A.M. & Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry* 33, 6739–6749.
37. Olah, G.A., Mitchell, R.D., Sosnick, T.R., Walsh, D.A. & Trewhella, J. (1993). Solution structure of the cAMP-dependent protein kinase catalytic subunit and its contraction upon binding the protein kinase inhibitor peptide. *Biochemistry* 32, 3649–3657.
38. Azevedo, W.F.D., Mueller-Dieckmann, H.-J., Schulze-Gahmen, U., Worland, P.J., Sausville, E. & Kim, S.-H. (1996). Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. *Proc. Natl. Acad. Sci. USA* 93, 2735–2740.
39. Xu, R.-M., Carmel, G., Kuret, J. & Cheng, X. (1996). Structural basis for selectivity of the isoquinoline sulfonamide family of protein kinase inhibitors. *Proc. Natl. Acad. Sci. USA* 93, 6308–6313.
40. Engh, R.A., Girod, A., Kinzel, V., Huber, R. & Bossemeyer, D. (1996). Crystal structures of catalytic subunit of cAMP-dependent protein kinase in complex with isoquinolinesulfonyl protein kinase inhibitors H7, H8, and H89. *J. Biol. Chem*. 271, 26157–26164.
41. Buechler, J.A. & Taylor, S.S. (1990). Differential labeling of the catalytic subunit of cAMP-dependent protein kinase with a water-soluble carbodiimide: identification of carboxyl groups protected by MgATP and inhibitor peptides. *Biochemistry* 29, 1937–1943.
42. Hemmer, W., McGlone, M., Tsigelny, I. & Taylor, S.S. (1997). Role of the glycine triad in the ATP-binding site of cAMP-dependent protein kinase. *J. Biol. Chem*., in press.
43. Goldberg, J., Nairn, A.C. & Kuriyan, J. (1996). Structural basis for the autoinhibition of calcium/calmodulin-dependent protein kinase I. *Cell* 84, 875–887.
44. Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J. & Pavletich, N.P. (1996). Crystal structure of the $p27^{Kip1}$ cyclin-dependent-kinase inhibitor bound to the cyclin A–Cdk2 complex. *Nature* 382, 325–331.
45. Zheng, J., Knighton, D.R., Parello, J., Taylor, S.S. & Sowadski, J.M. (1991). Crystallization of catalytic subunit of adenosine cyclic monophosphate-dependent protein kinase. *Methods Enzymol*. 200, 508–521.
46. Hamlin, R. (1985). Multiwire area X-ray diffractometers. *Methods Enzymol*. 114, 416–452.
47. Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol*. 33, 491–497.
48. Howard, A.J., Nielsen, C. & Xuong, N.-h. (1985). Software for a diffractometer with multiwire area detector. *Methods Enzymol*. 114 A, 452–472.
49. Brünger, A.T. (1990). Extension of molecular replacement: a new search strategy based on patterson correlation refinement. *Acta Cryst. A* 46, 46–57.
50. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* 235, 458–460.
51. Tronrud, D.E., Ten Eyck, L.F. & Matthews, B.W. (1987). An efficient general purpose least-squares refinement program for macromolecular structures. *Acta Cryst. A* 43, 489–501.
52. Cambillau, C. & Horjales, E. (1987). TOM: a FRODO subpackage for protein-ligand fitting with interactive energy minimization. *J. Mol. Graph*. 5, 174–177.
53. McRee, D.E. (1992). A visual protein crystallographic software system for X11/Xview. *J. Mol. Graph*. 10, 44–46.
54. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst*. 26, 283–291.
55. Hooft, R.W.W., Vriend, G., Sander, C. & Abola, E.E. (1996). Errors in protein structures. *Nature* 381, 272.
56. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355, 472–475.
57. Ramachandran, G.N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem*. 23, 283–438.

58. Venkatachalam, C.M. (1968). Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers* 6, 1425–1436.
59. Chou, P.Y. & Fasman, G.D. (1977). β-turns in proteins. *J. Mol. Biol*. 115, 135–175.
60. Milner-White, E.J., Ross, B.M., Ismail, R., Belhadj-Mastefa, K. & Poet, R. (1988). One type of gamma turn, rather than the other, gives rise to chain reversal in proteins. *J. Mol. Biol*. 204, 777–782.
61. *Insight II*, Version 95.0.4. (1995). Biosym Technologies/MSI, Inc., San Diego, USA.
62. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. *Biophys. J*. 64, 166–170.
63. Cheng, H.C., *et al.*, & Walsh, D.A. (1986). A potent synthetic peptide inhibitor of the cAMP-dependent protein kinase. *J. Biol. Chem*. 261, 989–992.